BECKER • KURIG • STRAUS
MÜNCHEN – BERLIN

Galileo®

Patentanwälte Becker Kurig Straus • Bavariastr. 7 • D-80336 München

Patent- och registreringsverket
Box 5055

S-102 42 Stockholm

SWEDEN

VIA TELEFAX IN ADVANCE
ORIGINAL BY REGISTERED MAIL

Patent- und Rechtsanwälte
European Patent Attorneys
European Trademark Attorneys

Dr. Eberhard Becker, Chem.
Dr. Thomas Kurig, Dipl.-Phys.
Dr. Alexander Straus, Dipl.-Chem.
Dr. Roman Vuille[1], Dipl.-Chem.
Friedrich von Braun, Rechtsanwalt

Peter Kylin, MSc.[2]
Magnus Hynell, MSc.[2]
Annika Björkman, MSc.[2]
Ivar Andréasson, MSc.[2]
Eva Lena Jansson[2]
Magnus Aspeby, MSc.[2] of counsel
Lars E. Johansson, MSc.[2] of counsel

Bavariastrasse 7
D-80336 München
Tel.: +49-89-746 303 0
Fax: +49-89-746 303 11
info@galileolaw.de
www.galileolaw.de

May 17, 2004

PCT-Application PCT/IB02/01228
Applicant / Owner: Nokia Corporation
Our Ref.: 50907 WO (KG/TP)

In Response to the First Written Opinion Dated February 23, 2004:

Enclosed please find

a new set of claims 1 to 27,

which substitute the claims being presently on file and on the basis of which further proceedings shall be carried out without prejustice.

I. New Claims

The new claim 1 is formed from the original claim 1 and the wording of the new claim 1 is clarified. In detail, the subject-matter as defined by the independent claim 1 is operable with hybrid speech coders and hybrid speech decoders set forth with reference to the original claim 22, page 9, lines 6 - 9 and page 25, lines 1 - 4 as well as claim 24, page 21 - 24 and page 25, lines 1 - 4 of the specification. The purpose of the derived phase-

characterizing parameter to prevent misalignment between the preceding and the subsequent frames is disclosed on page 4, lines 6 - 10 of the present spec.

In order to state the sequence of the frames in question defined in the independent claim 1 more precisely, the frame coded according to the waveform matching speech coding is designated as preceding frame, whereas the frame according to the parametric speech coding is designated as subsequent frame. A corresponding disclosure relating the preceding frame coded according to the waveform matching speech coding can be obtained from page 4, lines 1 - 2. Moreover, a disclosure relating to the subsequent frame coded according to the parametric speech coding can be gathered from the teaching on page 4, lines 6 - 10, which also discloses the immediate succession of the frames in question.

The new claims 2 to 11 correspond to the original claims 2 to 11.

The new independent claim 8 is adapted to the wording of the new independent claim 1.

The new claims 12 to 27 correspond to the original claims 12 to 27.

The numbering and back references of the new claims are adapted.

II. State of the Art

The Examiner refers to a state of the art document, which is

D1: Katugampala N. and Kondoz A., *"A hybrid coder based on a new phase model for synchronization between harmonic and waveform coded segments"*, Electronic Letters, vol. 2, pages 685 - 688, May 2001, ISBN 0-7803-7041-4

D1 is the closest prior art document relevant for the present invention.

D1 relates to a hybrid coder with a phase model to synchronize harmonic and waveform coded segments. The coder of D1 is based on an analysis by synthesis to distinguish between stationary and transitional segments. Harmonic excitation is synchronized with a linear (LPC) residual by transmitting the location of the pitch pulse closest to the frame boundary and a phase value that represents the shape of the corresponding pitch pulse.

III. Object

A first object of the present invention is to provide a method and a device capable for operating the method for determining phase characteristics of a frame coded according to a waveform matching coding for providing the determined phase characteristics as an initial phase characteristics for parametric coding of a succeeding speech frame.

A second object of the present invention is to provide a method and a device capable for operating the method for detecting misalignments between a preceding frame coded according to a waveform matching coding and a frame succeeding (immediately) the preceding frame coded according to a parametric speech coding to enable a re-coding of the succeeding frame.

IV. Solution

The first object of the present invention is solved by a method operable with hybrid speech coders and hybrid speech decoders, which method provides at least one phase-characterizing parameter. The phase-characterizing parameter is derived from characteristics obtained from a preceding waveform matching frame and applicable for a parametric frame, which succeeds in time to the waveform matching frame. The phase-characterizing parameter enables to prevent misalignment between the frames in question.

The second object of the present invention is solved by a method operable with a network device for ensuring enhanced speech quality, which method evaluates characteristics obtained from a preceding waveform matching frame and a succeeding parametric frame, which succeeds in time to the waveform matching frame, such that misalignment between the frames in question are detectable.

V. Novelty and Inventive Step

D1 establishes the closest prior art relevant in view of the present invention.

D1 teaches the synchronizing between harmonic and waveform coded segments, which is obtained by the means of the synchronized waveform matched phase model (SWPM) adapted to preserve time synchrony and waveform shape between original speech and synthesized speech. This methodology of D1 relates rather to a classical method to maintain synchrony between the synthesized, i.e. synthesized by a parametric speech

coder, and the original speech. In the fully voiced frames, where a parametric speech coder is employed, supplementary phase information is determined and transmitted to the speech decoder. This phase information is based on the pitch pulse location and pitch pulse shape in the original speech and in the voiced frame. In the succeeding voice frames the phase information is interpolated between the phase value at the end of the previous frame and the transmitted value for the current frame.

In contrast to D1, the present invention relates to transients between frames coded by a waveform coder and a parametric coder. To get the phase value in the first parametric frame, the estimated pitch pulse value in the waveform frame is exploited to ensure smooth transition. Consequently the synchrony between the original and coded speech needs not to be remained, which is asserted as a basic concept of the teaching in accordance with D1. The present invention reflects the requirement to guarantee a smooth transition between succeeding frames coded by waveform speech coding and parametric speech coding. How to handle this specific and very crucial transition in terms of quality is not discussed in D1 since the teaching thereof concentrates merely on the problem of phase continuation between the voiced frames.

Furthermore, technology relating to speech coding, handling of exceptions and damage control is considered very crucial for optimal performance. The present invention illustrates possible failures in consequence to which methods are presented to overcome and prevent these. One method for phase estimation is not optimal in all cases and failures such as missing pitch pulse may occur.

The basis concept of the present invention concerns methodologies and realizations where the pitch continuity is checked during the transitional regions. D1 does not include any mechanism for failure prevention or exception handling, but instead uses always the same methodology regardless whether it is successful or not.

Conclusively, the methods, devices and system defined in the claims accompanying the present invention are not anticipated by the closest prior art document D1.

As aforementioned, the closest prior art document D1 provides a closed teaching which relates to the phase continuation between the voiced frames and solves the problem by transmitting supplementary phase information being based on the pitch pulse location and pitch pulse shape in the original speech and in the voiced frame. The phase information of D1 enables to ensure the synchrony between original and synthesized speech. Due to the

fact that D1 defines a closed teaching, the inventive concept and inventive teaching according to the present invention in neither included in D1 nor suggested by D1. This fact is also visible to those skilled in the art when referring to the concept according to the present invention to detect and/or eliminate misalignments between frames in question at a network device disposed intermediately within the transmission link of terminals communicating said frames.

Consequently, the claims being presently on file of the invention are likewise assumed as inventive over the closest state of the art document D1.

VI. Requests

The Examiner is kindly requested to reconsider its opinion about the patentability of the present invention in view of the arguments provided above and with respect to the accompanying set of new claims.

Dr. Thomas Kurig
(Patent Attorney)

Enclosure: Set of new claims 1 to 27

PCT-Application PCT/IB02/01228
Applicant: Nokia Corporation
Our Ref.: 50907 WO WO (KG/TP)

New Claims

1. Method for providing at least one phase-characterizing parameter for speech processing operable with hybrid speech coders and hybrid speech decoders, comprising:
   - obtaining characteristics of a preceding frame coded according to a waveform matching speech coding; said preceding frame according to said waveform matching speech coding being immediately preceding in time to a succeeding frame according to a parametric speech coding

   characterized by
   - deriving said at least one phase-characterizing parameter for processing said succeeding frame according to said parametric speech coding from said obtained characteristics;

   wherein said at least one phase-characterizing parameter is employable to prevent a misalignment of said frames.

2. Method according to claim 1, wherein said speech processing is a speech encoding operation.

3. Method according to claim 1, wherein said speech processing is a speech decoding operation.

4. Method according anyone of the preceding claims, wherein said step of obtaining characteristics of said preceding frame according to said waveform matching speech coding comprises:
   - determining positions of at least one pulse of said preceding frame according to said waveform matching speech coding; and
   - determining a position of a last pulse of said at least one pulse.

5. Method according to claim 4, wherein said at least one pulse is at least one pitch pulse.

6. Method according to claim 4 or claim 5, wherein said step of obtaining characteristics of said preceding frame according to a waveform matching speech coding comprises:
   - determining a pulse value from the distances between said at least two pulses.

7. Method according to claim 4 or claim 5, wherein said obtaining characteristics of said preceding frame according to a waveform matching speech coding comprises:
   - obtaining a pulse value from an antecedent frame.

8. Method according to claim 6 or claim 7, wherein said at least one phase-characterizing parameter is obtained from said position of said last pulse relative to a size of said preceding frame according to said waveform matching speech coding in relation to said pulse value.

9. Method according to anyone of the preceding claims, wherein said at least one phase-characterizing parameter is at least one phase value.

10. Method according to anyone of the claims 2 to 9, wherein said determining of said positions comprises:
- determining average energy values from said preceding frame according to said waveform matching speech coding and
- evaluating said average energy values in order to determine positions of at least one local maximal energy value and
- assigning said positions of said at least one local maximal energy value to said positions of said at least one pulse.

11. Method according to claim 10, wherein said determining said average energy values comprises the step of:
- employing a sliding average algorithm in order to determine said average energy values.

12. Method for detecting a transition misalignment in transition from a preceding frame according to a waveform matching speech coding to a succeeding frame according to a parametric speech coding, said preceding frame according to said waveform matching speech coding being immediately preceding in time to said succeeding frame according to said parametric speech coding, comprising:
- obtaining characteristics of said preceding frame according to said waveform matching speech coding,
- obtaining characteristics of said succeeding frame according to said parametric speech coding, and
- evaluating said obtained characteristics in order to detect said transition misalignment.

13. Method according to claim 12, wherein said obtaining characteristics of said preceding frame according to said waveform matching speech coding comprises:
- determining positions of at least one pulse from said preceding frame according to said waveform matching speech coding and
- determining a position of a last pulse of said at least one pulse,

and wherein said obtaining characteristics of said succeeding frame according to said parametric speech coding comprises:

- determining positions of at least one pulse from said succeedingframe according to said parametric speech coding and
- determining a position of a first pulse of said at least one pulse,

14. Method according to claim 13, wherein said pulses are pitch pulses.

15. Method according to claim 13 or claim 14, wherein said evaluating said obtained information comprises:

- determining a distance of said position of said last pulse and said position of said first pulse and
- comparing said distance with a pulse value.

16. Method according to claim 15, wherein said pulse is obtained by the step of:

- determining said pulse value from distances of said pulses included in said preceding frame according to said waveform matching speech coding.

17. Method according to claim 15, wherein said pulse is obtained by the step of:

- determining said pulse value from a phase contour of an antecedent frame according to said parametric speech coding.

18. Method according to anyone of the claims 12 to 17, wherein said determining of said positions comprises:

- determining average energy values from said frame and
- evaluating said average energy values in order to determine positions of at least one local maximal energy value and
- assigning said positions of said at least one local maximal energy value to said positions of said at least one pulse.

19. Software tool for speech processing, comprising program code portions for carrying out the operations of any one of claims 1 to 18, when said program is implemented in a computer program for executing on a computer, a user terminal or a network device.

20. Computer program for speech processing, comprising program code section for carrying out the operations of any one of claims 1 to 18, when said program is run on a computer, a user terminal or a network device.

21. Computer program product for speech processing, wherein said computer program product is comprising program code sections stored on a computer readable medium for carrying out the method of any one of claims 1 to 18, when said program product is run on a computer, a user terminal or network device.

22. Communication terminal device offering enhanced quality of transmitted speech data comprising a speech encoder including a parametric speech encoding unit, a waveform matching speech encoding unit, and a communication interface for communicating speech encoded data via a mobile communication network, wherein said speech encoder is able to operate the method for providing at least one phase-characterizing parameter for coding a succeeding frame according to a parametric speech coding according to anyone of the claims 1 to 11.

23. Communication terminal device offering enhanced quality of transmitted speech data comprising a speech decoder including a parametric speech decoding unit and a waveform matching speech decoding unit and a communication interface for communicating speech encoded data via a mobile communication network, wherein said speech decoder is able to operate the method for detecting a transition misalignment in transition from a preceding frame according to a waveform matching speech coding to a succeeding frame according to a parametric speech coding according to anyone of the claims 12 to 18.

24. Terminal device according to claim 23, said speech decoder being additionally able to operate the method for providing at least one phase-characterizing parameter for coding a succeeding frame according to a parametric speech coding according to anyone of the claims 1 to 11.

25. Network device offering enhanced quality of transmitted speech data comprising a communication interface for receiving encoded speech data and transmitting encoded speech data and an analyzing unit, said analyzing unit being able to operate the method for detecting a transition misalignment from a preceding frame according to a waveform matching speech coding to a succeeding frame according to a parametric speech coding according to anyone of the claims 12 to 18.

26. Network device according to claim 22, said analyzing unit being additionally able to operate the method for providing at least one phase-characterizing parameter for coding a succeeding frame according to a parametric speech coding according to anyone of the claims 1 to 11.

27. System offering enhanced quality of transmitted speech data comprising:
    - a first terminal comprising a speech encoder for encoding speech and a communication interface for transmitting encoded speech data,
    - a first terminal comprising a speech decoder for decoding said encoded speech data and a communication interface for receiving said encoded speech data,
    - an intermediate network device offering enhanced quality of transmitted speech data according the anyone of the claims 25 to 26.